Astrobiology. Volume 17, Issue 10, October 2017, pp. 1009-1021
DOI:10.1089/ast.2016.1575

Correlations between life-detection techniques and implications for sampling site selection in planetary analogue missions

Diana M. Gentry[a,*], Elena S. Amador[b], Morgan L. Cable[c], Nosheen Chaudry[d], Thomas Cullen[d], Malene B. Jacobsen, Gayathri Murukesan[e], Edward W. Schwieterman[b], Adam H. Stevens[f], Amanda Stockton[g], George Tan[g], Chang Yin[h], David C. Cullen[d], and Wolf Geppert[h]

[a]Biospheric Science, NASA Ames Research Center, Moffett Field, CA 94035, USA

[b]Astrobiology Program, University of Washington, 4000 15th Ave NE, Seattle, WA 98195, USA

[c]NASA Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, USA

[d]School of Engineering, Cranfield University, College Road, Cranfield, MK43 0AL, Bedfordshire, UK

[e]University of Turku, Department of Biochemistry/Biochemistry, 20014 Turun Yliopisto, Finland

[f]UK Centre for Astrobiology, School of Physics and Astronomy, The University of Edinburgh, EH9 3FD, Edinburgh, UK

[g]School of Chemistry & Biochemistry, Georgia Institute of Technology, 901 Atlantic Drive, Atlanta, GA 30332, USA

[h]Astrobiology Centre, Stockholm University, SE - 106 91, Stockholm, Sweden

[*]To whom correspondence should be addressed: NASA Ames Research Center, MS 245-4, Moffett Field, CA 94035, USA, +1 650 604 5441, diana.gentry@nasa.gov

Abstract

We conducted an analogue sampling expedition under simulated mission constraints to areas dominated by basaltic tephra of the Eldfell and Fimmvörðuháls lava fields (Iceland). Sites were selected to be 'homogeneous' at a coarse remote sensing resolution (10 – 100 m) in apparent color, morphology, moisture, and grain size, with best-effort realism in numbers of locations and replicates. Three different biomarker assays (counting of nucleic-acid-stained cells via fluorescent microscopy, a luciferin/luciferase assay for adenosine triphosphate (ATP), and

quantitative polymerase chain reaction (qPCR) to detect DNA associated with bacteria, archaea, and fungi) were characterized at four nested spatial scales (1 m, 10 m, 100 m, and > 1 km) using five common metrics for sample site representativeness (sample mean variance, group *F*-tests, pairwise *t*-tests, and the distribution-free rank sum *H* and *u*-tests). Correlations between all assays were characterized with Spearman's rank test. The bioluminescence assay showed the most variance across the sites, followed by qPCR for bacterial and archaeal DNA; these results could not be considered representative at the finest resolution tested (1 m). Cell concentration and fungal DNA also had significant local variation, but were homogeneous over scales of > 1 km. These results show that the selection of life detection assays and the number, distribution, and location of sampling sites in a low biomass environment with limited *a priori* characterization can yield both contrasting and complementary results, and that their interdependence must be given due consideration to maximize science return in future biomarker sampling expeditions.

Keywords: astrobiology, biodiversity, microbiology, Iceland, planetary exploration, Mars mission simulation, biomarker

Introduction

Investigation of potential extraterrestrial habitats and biomarkers relies on a combination of infrequent *in situ* robotic planetary exploration missions and more

readily available, but far lower resolution, remote sensing data. A few extraterrestrial sample return missions have been carried out (Stardust, Hayabusa, Apollo), and others are in progress or planned (Hayabusa 2, OSIRIS-REx, Chang'e 5), including multi-stage missions led by sample collection and caching rovers (Mars 2020). However, the expense, technical challenges, and planetary protection requirements of such missions means the number of samples returned will remain limited for the foreseeable future. It is therefore necessary to understand how sample site selection in such constrained mission contexts impacts scientific return. In a life-detection mission, where many assays require consumables, this requirement is even more critical.

When the objective is to gain information about a planetary surface, determining the representativeness of a potential sample set is a significant challenge. The difficulty is compounded still further by limited *a priori* knowledge about what can be very complex climatic, mineralogical, and chemical macro- and micro-environments. Mars is currently the best-case scenario for characterization of potential sample return sites, which may include up to 30 cm/pixel orbital imaging (e.g., HiRISE) or 10 μm to 1 mm scale *in situ* measurements (*e.g.*, MAHLI, ChemCam) if precursor rover or lander missions are included. However, both the number and diversity of potential target environments for planetary exploration could be significantly increased with improved use and contextualization of lower-

cost, faster-turnaround planetary data (*e.g.*, 100 m scale flyby imaging or 10 cm scale descender/lander measurements). In such a situation, analogue work is the best available guide to understanding how the target biomarker(s) might be expected to vary not just over different environments, but also within apparently homogeneous environments. Such an understanding is particularly beneficial to missions whose short lifetime or need for consumable reagents further reduce the potential number of analyses. For example, the Mars Science Laboratory Sample Analysis at Mars (SAM) instrument suite carries only nine wet chemical derivatization cells and 54 pristine pyrolysis chambers (Mahaffy et al., 2012); the CheMin instrument carries 74 pristine sample containers (Blake et al., 2012); the proposed fluidics-based Signs of LIfe Detector (SOLID) holds 16 microarray chips (McKay et al., 2013); and the proposed Life Marker Chip, also microarray-based, is designed for as few as four samples (Sims et al., 2012).

The need to understand the likely distribution and variability of potential biomarkers in extraterrestrial environments highlights the need for a better understanding of these distributions on Earth, particularly in planetary analogue environments where biomass is very low. Existing studies have shown that, in general, the spatial dependence of terrestrial microbial abundance and diversity is extremely complex (see Green and Bohannan (2006) for a review). It is typically strongly affected by the local scale of environmental variation; in the same location,

different types of organisms may vary on different spatial scales over less than a centimeter in some cases to over hundreds of kilometers in others (Pasternak et al., 2013), and each class may be organized on multiple spatial scales simultaneously (Franklin and Mills, 2003).

Much work on life detection and biomarker characterization has been done in extreme environments on Earth, such as the cold, dry valleys of Antarctica or the ancient, high-altitude Atacama desert, that are considered to be the closest available analogues to regions of Mars (Preston and Dartnell, 2014). For studies of potential habitats on Mars, Icelandic lava fields are often chosen due to their volcanic geochemistry, interactions with glaciers, range of recent ages, and relative isolation from anthropogenic activity (*e.g.*, Warner and Farmer (2010); Bagshaw et al. (2011); Cousins and Crawford (2011)).

Though a significant amount of work has been done in landscape ecology and related fields characterizing terrestrial soil microbial biodiversity (Peigné et al., 2009), many of the studied parameters are not relevant to planetary exploration or this type of analogue environment (*e.g.* effects of land use, tree and bush root effects (Ettema and Wardle, 2002)). Further, the generalizability between results from these biomass-rich environments (*e.g.* Naveed et al. (2016)) and the low-biomass, unusually nutrient-limited environments used as planetary analogues is uncertain. For example, there is some evidence that environments in which

overall microbial abundance is lower have more spatially- and temporally-diverse communities (*e.g.*, Banerjee et al. (2011) and references reviewed in Kieft (2003)).

Given the lack of studies directly addressing variability and correlation among biomarker types in Mars analogue environments, this simulated mission is an initial effort using statistical analyses in common use in the astrobiology community and which directly quantify differences between groups of samples. Future planned expeditions are encouraged to include additional physicochemical parameters, integration of remote sensing data, and more specific geostatistical techniques such as autocorrelation to allow development of a model of spatial dependence of biomarker types in this type of analogue environment.

We conducted a planetary exploration analogue mission to multiple Icelandic lava fields in July 2013 with several interrelated goals. Although the sites chosen are Mars analogue environments, and thus the operational constraints and specific data are most relevant to future Mars missions, assays were chosen to represent broad categories of potential biomarkers (biomass, metabolic activity, species-specific indicators) also relevant to other potential planetary targets. In a companion paper (Amador et al., 2014), we described the feasibility of using three different life detection techniques in multiple cycles of on-site sampling, analysis, and follow-up site selection under simulated robotic constraints in numbers of samples and assays performed; these constraints are briefly summarized in the Methodology

section. Here, we report the results of the life detection assays and their variation at different spatial scales across sites chosen to be 'homogeneous' at the level of typical low-resolution (10 – 100 m) orbital remote sensing data in apparent color, morphology, moisture, and grain size.

Cell staining and fluorescence microscopy was used to observe and count cells extracted from the tephra matrix. Adenosine triphosphate (ATP) bioluminescence was used to quantify metabolic activity and bioavailable energy. Quantitative real-time polymerase chain reaction (qPCR) assays were used to quantify distribution of archaea, bacteria, and eukaryotes. Several common statistical tests used to examine differences between groups of samples – pairwise mean and median comparisons (Welch *t*- and Mann-Whitney *u*-tests with Bonferroni correction) and parametric and non-parametric analysis of group variance (ANOVA's *F*- and Kruskal-Wallis *H*-tests) – were performed for the results of each assay to determine their typical variation at different scales, and correlation coefficients (Spearman ρ) were calculated for each pair of assays to determine how the assays' results complemented each other.

Methodology

Sampling methodology

Samples were collected from two geologically recent lava fields: Fimmvörðuháls (2010 eruption, 63° 38' 12.30" N, 19° 26' 49.20" W) and Eldfell on Heimaey,

Westman Islands (1973 eruption, 63° 25' 08.30" N, 20° 14' 38.70" W). Separated by 45 km, these fields have broadly similar alkali basalt compositions (Higgins and Roberge, 2007; Sigmundsson et al., 2010), ranges of sediment types, and limited vegetation. Sampling locations were chosen in nested grids of 1 m, 10 m, and 100 m resolution (FIG. 1) at both sites.

Fimmvörðuháls samples ranging from coarse ash (grain size < 2 mm) to lapilli (grain size 2 – 64 mm) were collected around the Magni cinder cone. Eldfell samples (also lapilli-sized or smaller) were collected around a large scoria cone. All samples were taken from ~5 cm depth at locations where tephra appeared to be homogeneous in visible color, morphology, moisture, and grain size (FIG. 2). Triplicate samples were taken from each location. Full sampling procedures, including precautionary measures to avoid anthropogenic contamination, are given in the companion paper. Two positive controls were collected as well: ash from the base of a small tuft of grass in Fimmvörðuháls, chosen to match the samples' overall geochemistry, and a soil sample from the grassy area around the field laboratory, chosen to ensure a very large microbial population. Sterile water was used as a negative control where required.

Mission and field constraints

Simulated mission constraints and practical field considerations are discussed in detail in Amador et al. (2014). Briefly, all sample sites were within one day's travel

of a field laboratory established in a nearby school, Hvolsskóli, in the town of Hvolsvöllur. The setup included access to municipal water, electricity, a teaching laboratory space and smaller classrooms, and a kitchen (FIG. 3). All instruments and consumables used in analyses were shipped by participants (*e.g.*, benchtop luminometer) within a single 119 cm × 79 cm × 52 cm aluminum trunk, although some support equipment was procured on site (*e.g.*, pressure cooker used as autoclave). All consumables chosen had to be stable during shipping and storage in an ordinary kitchen freezer. A total of 45 sample sites were assayed, with no more than 3 samples taken per 1 m site. Sample analysis was required to be completed within a two-day turnaround, and all assays shared upstream preparation (grinding and homogenization). Throughput limited the fluorescence microscopy assay to a single replicate per sample and the qPCR assays to two replicates for most samples.

Life detection assays

Three established life detection techniques were chosen: sample washing followed by cell staining with fluorescence microscopy (Nadeau et al., 2008); sample washing and cell lysis followed by an adenosine triphosphate (ATP)-detecting luciferin/luciferase luminescence assay (Obousy et al., 2000; Fajardo-Cavazos et al., 2008); and quantitative polymerase chain reaction, abbreviated qPCR (Fajardo-Cavazos et al., 2010). The three were selected from among many initially suggested techniques due to their rapid turnaround time and suitability for implementation in a field lab (portable instrumentation and low cost, easy-to-store

reagents).

Fluorescence microscopy, a method of direct cell quantification, was performed here using a commercial stain that binds to double-stranded DNA. When cells in liquid suspension are stained and placed under appropriate excitation and emission filtering, the concentration of DNA within intact cells (living or dead) will cause them to appear as bright objects against a dark background. At the appropriate dilution, individual cells can be counted by eye or with image-processing software, allowing the microbial abundance of a sample to be quantified. Cell quantification by fluorescence staining can be implemented with many commercially available stains using a benchtop microscope with standardized filter sets and illumination.

Adenosine triphosphate (ATP) is a general indicator of bioavailable energy, stable under many 'life-friendly' environmental conditions, and easy to detect, making it a widely studied and prioritized biomarker (Parnell et al., 2007). The ATP bioluminescence assay relies on the enzyme luciferase's consumption of ATP during conversion of luciferin to an excited state of oxyluciferin, which in then decays under photon emission; when luciferin and luciferase are added to the product of cell lysis, an amount of light is generated proportional to the amount of ATP liberated from the cells. A greater concentration of ATP generally indicates cells with higher metabolic activity, although the relationship with cell concentration or biomass is complicated by both the need for cell lysis, which is

not equally efficient for all cell types, and by inherent differences in cell ATP content due to cell state (*e.g.*, sporulating, dividing), level of metabolism, species, and other factors (Stanley, 1989). The ATP bioluminescence assay can be implemented with commercially available pre-mixed reagents and a benchtop luminometer.

The quantitative polymerase chain reaction (qPCR) assay allows for simultaneous amplification and quantification of DNA that matches the set of primers used. It combines the standard PCR amplification process with a dye that is activated as it binds to DNA. By choosing an array of primers that correspond to different taxonomic groups (*e.g.*, bacteria, archaea, or fungi) and using each set on a different sample aliquot, it is possible to assess both the quantity of DNA present in a given sample and its relative diversity (defined for our purposes as the relative ratios of recovered DNA from each group). Although bacterial, archaeal, and fungal DNA markers are Earth-specific, they correspond to types of single-celled life likely to have different dispersion methods and habitat preferences, and in an analogue context can be viewed as simply 'different' primer sets designed to cover a wide range of potential life. This implementation of qPCR requires pre-synthesized primers, commercially available buffer solutions, and a thermocycler with an integrated well-plate fluorometer.

Sample preparation and analysis

50 mL tephra samples were taken from each site as described above and stored at room temperature in the field laboratory. To break up large particulates, each sample was placed in layered, sterile plastic bags (Whirl-Pak®;), the bags wrapped in a clean laboratory towel (Technicloth®;), and the wrapped bags crushed using a hammer or table vise grip sterilized with 70% isopropanol until bulk grain size appeared ≤ 2 mm. After this initial common preparation step, the samples were divided into 1 mL portions and passed into the sample preparation pipelines for individual techniques.

Fluorescence microscopy

After initial sample preparation as described above, each 1 mL sample portion dedicated to the fluorescence microscopy procedure was loaded in a 1.5 mL microcentrifuge tube with 0.75 mL of sterile PBS/Tween buffer, inverted by hand until the sample was fully wetted, and then mixed with a vortexer. The tube was then placed in an ultrasonic bath for 5 minutes to detach any remaining organisms from particulate surfaces. After sonication, the sample was gently centrifuged for 5 minutes at 600×*g*, causing suspended sample particulates to sediment out while leaving any microbes in suspension. The supernatant was then drawn off into a separate sterile tube, and the remaining sediment resuspended in PBS/Tween and subject to the same repeated process.

The supernatants from both cycles were then combined, mixed, loaded into a sterile

2 mL syringe, and filtered through a Whatman Nuclepore Track-Etch 0.2 μm membrane filter. The syringe was refilled twice with sterile water and passed through the same filter, and twice more with air. The filter containing the concentrated sample filtrand was then removed and placed on a slide for staining. To enable counting of all cells, live or dead, 100 μL of SYBR Gold Nucleic Acid Gel Stain (Molecular Probes, Inc.) was spread evenly over the area of the filter. A cover slip was applied and the stained sample incubated in the dark, at room temperature, for at least 15 minutes before imaging.

The counting protocol largely followed that recommended in Kepner and Pratt (1994), with adaptations to account for the use of a digital imager. The stained filter membranes were viewed under a Partec CyScope (fluorescence and transmitted light microscope) equipped with a 455 nm emission light source and 500 nm dichroic mirror (DM) long pass filter. Each filter, corresponding to one extracted and filtered subsample, was imaged in five randomly chosen locations using the 100× (oil) objective and the stained cells in each location manually counted. Digital micrographs of each field of view were also taken for later verification and analyses.

The mean cell counts per field of view were converted to an estimate of cell concentration per 1 mL of sample by dividing by the fraction of the filter surface area counted and the fraction of wash suspension present on the filter. As this

assumes 100% wash efficiency and filtration, it yields a conservative estimate of total microbial abundance.

ATP bioluminescence

After initial sample preparation as described above, each 1 mL sample portion dedicated to ATP analysis was washed with sterile Tris-EDTA buffer (100 mM Tris, 4 mM EDTA) and filtered to remove sediment. The resulting suspensions were lysed by incubation in boiling water (100 $^{\circ}$C) for 2 minutes. The lysis products were then further divided into four 100 μL aliquots.

The amount of ATP present in each aliquot was determined using the Roche ATP Bioluminescence Assay Kit HS II, following the manufacturer's protocol for a tube assay with the modifications described in Barnett (2010). All standards and samples were cooled to 4 $^{\circ}$C before measurement by placing them on ice. The standards and samples were read by adding 100 μL of luciferase reagent to each tube and reading the resulting luminescence with a Merck HY-LiTE 2 portable luminometer.

The luminometer yields a measurement in relative light units (RLU). To translate these relative values into absolute ATP quantification, a calibration curve of RLU readings (R) from known molar concentrations of ATP (C_{ATP}) was generated on site in the field laboratory using the same protocol and reagents. The best-fit linear relationship established by the calibration curve was used to convert the sample

RLU data into estimated ATP molar concentration.

Quantitative PCR

After initial sample preparation as described above, each 1 mL sample portion dedicated to qPCR analysis was subjected to DNA extraction and purification using a PowerSoil®;DNA Isolation Kit (MO BIO Laboratories, Inc.) following the manufacturer's protocol. The purified DNA suspension was diluted 1:10 to create an appropriate starting concentration. From this suspension, two duplicate 8 μL aliquots were taken and each combined with 10 μL of SYBR Green Supermix (BioRad) and 1 μL each of forward and reverse oligonucleotide primer, for a total of 20 μL each in two wells of a 96-well plate. qPCR was performed in duplicate for each full plate using a BioRad MiniOpticon^{TM} real-time PCR system.

As the total number of plates that could be run in the field laboratory was limited by time constraints and reagent availability, a small set of primers was selected (TABLE 1) to target the 16S/18S region across a high-level taxonomic range (bacteria, archaea, or fungi). All primers were also chosen to have the same extension temperature to ensure that all plate wells on each run could be filled. The extension temperature for all three primers and runs was 54 $^{\circ}$ C.

Levels of the target 16S/18S genes in the total extracted DNA for each sample were calculated using the standard $2^{-\Delta Cq'}$ method (Livak and Schmittgen, 2001) (assuming 100% amplification efficiency). As the constraints of the field lab setup

did not allow separate calibration plates to be run, only relative initial concentrations (scaled to the minimum measurable value) of each amplicon could be determined.

Results and Analysis

Lessons Learned

We used rapid, in-field analysis to determine subsequent sampling decisions, as described in detail in our prior companion publication (Amador et al., 2014). Briefly, samples were collected by a field expedition team while simultaneously the field laboratory team analyzed samples collected on the prior day's expedition. We found that the fluorescence microscopy protocol used for this work was too slow for meaningful sample throughput, with the combination of limited field of view and sparse filtrate leading to low means and high variances in the cell count data. Manual cell counts also introduce a high potential for human error and, in a field lab context, reduced capacity for replicates. Future work could be improved by re-suspending and concentrating the cells post-filtration and by the use of assistive imaging technology, such as ImageJ (Schneider et al., 2012). The ATP assay was well-suited to our approach, as was the qPCR assay, especially if a higher-throughput qPCR instrument were to be used.

Fluorescence microscopy

The mean cell concentrations for each sample site, at all four nested spatial scales, are shown in FIG. 4 and FIG. 5. Histograms at each scale are shown in FIG. S1,

FIG. S2, and FIG. S3. Unlike the ATP and qPCR assays, only a single replicate from each 1 m site was quantified with fluorescence microscopy due to operational constraints; the values for the 1 m sites are arithmetic means of five randomly chosen fields of view for a single filter. Overall, cell count values were fairly low. The data was analyzed following procedures detailed in Mendenhall and Sincich (1995).

To quantify how much variability within sites was present, the unbiased sample variance s^2 of each sample grouping was calculated. The mean variance for samples at each spatial scale is shown in TABLE 2. The mean variances in cell concentration within sites do not follow a trend with spatial scale, indicating that differences in cell concentration between sites do not significantly change with distance. Overall, the variances for the cell concentration data are high in comparison to the sample group means.

One common measure of sample site representativeness is the degree to which the differences between sites exceed differences within sites across all sites in a given group. Probably the most familiar metric for this, due to its use in ANOVA, is the *F* statistic: the lower the resulting *p*-value, the more significant the difference in results between sites grouped at that scale (or, equivalently, the greater the effect of sample site choice at that scale on mean result). The *F*-test results are shown in TABLE S1; some diversity is apparent at the 10 m ($p = 0.03$) and 100 m ($p = 0.003$)

scales, but not at the > 1 km scale ($p = 0.96$). In response to the observation below about the outlier status of FIM–6–1, those three 1 m sites were removed from the data set and the *F*-tests rerun; in these results, the 10 m results are no longer significant ($p = 0.30$), and the 100 m results are borderline ($p = 0.09$). Regardless of whether the potential outlier site FIM–6–1 is included, the cell count assay data shows far less variation between sites than the ATP or qPCR assay data.

A complementary test indicating how many individual sites within a scale are different and by how much, two-tailed *t*-tests using Bonferroni's multiple comparisons procedure, was conducted for differences between sampling site mean cell concentrations between each pair of sampling sites at each spatial scale. Only a single 10 m site pair, FIM–1–1 and FIM–6–1, are significantly different at $\alpha = 0.05$, equivalent to 1% of 10 m site pairs. At the 100 m scale, three pairs of sites (10%) are different, HEI–2 and FIM–6, FIM–1 and FIM–6, and FIM–4 and FIM–6. All four appear to result from the unusually high readings at FIM–6–1; this sampling site was observed to have a noticeably high fraction of reddish particles (FIG. 6). This may indicate the presence of more oxidized iron, which in turn may indicate a longer history of moisture exposure. At the > 1 km scale, the two field sites are not significantly different.

In robotic exploration contexts, the number of replicates taken from a given sample site is likely to be very small; for instance, the CheMin instrument has analyzed as

few as one drill sample per km at times (Treiman et al., 2016). Both the *t*- and *F*-tests may give misleading results in such cases if the underlying distribution of data – in our case, of cell concentration between and within sites – is not a normal distribution. To account for this possibility, the analyses were repeated using the non-parametric equivalent *u*- and *H*-tests (TABLE S2), which require no assumptions about distribution and and test for changes in sample data such as shifted medians. The *H*-tests indicate borderline-significant diversity at the 10 m (p = 0.05) and 100 m (p = 0.02) scale; repeated without the FIM–6–1 samples, the significance is reduced somewhat, to p = 0.11 at 10 m and p = 0.06 at 100 m. The *u*-tests, due to the reduced number of replicates in this data set, cannot yield significant results at the 10 or 100 m scale, but do confirm the lack of difference between Eldfell and Fimmvörðuháls.

ATP bioluminescence

The mean ATP concentrations for each sample site, at all four nested spatial scales, are shown in FIG. 4 and FIG. 5. Histograms at each scale are shown in FIG. S1, FIG. S2, and FIG. S3. Overall, samples from Eldfell showed noticeably higher concentrations, though both were far lower than the levels of ATP in the positive control samples taken from areas with visible vegetation. The data was analyzed as described below following the same procedure as the fluorescence microscopy data.

For mean ATP concentration, sample variance increases with scale, indicating (matching intuition) that sites which are separated by larger distances tend to be more different from each other. The *F*-tests show that all spatial scale groupings (1 m, 10 m, 100 m, and > 1 km) have strong effects ($p < 0.001$), indicating that there is, overall, large and significant diversity in ATP concentration at all spatial scales examined. The pairwise *t*-tests show that at the 1 m scale, approximately 45% of sites are different from each other at a significance level of $\alpha = 0.05$, 41% at 10 m, and 29% at 100; the two > 1 km sites are also significantly different (full results given in supplementary tables).

As with the fluorescence microscopy data, non-parametric *H*- and *u*-tests were conducted. The *H*-tests, consistent with the *F*-tests, show significant diversity in relative ATP concentration at all spatial scales examined. At the 1 m scale, the combination of few replicates (4) and large number of comparisons (45 sites = 990 pairs) does not allow conclusions to be drawn at any significance level from pairwise *u*-tests; however, at the 10 m scale and 100 m scales, 50% of sites are different from each other at a significance level of $\alpha = 0.05$, and the two > 1 km sites are again significantly different.

Quantitative PCR

The relative mean genomic DNA concentrations (scaled to the minimum measured value), for each sample site, at all four nested spatial scales, are shown in FIG. 4

and FIG. 5. Histograms at each scale are shown in FIG. S1, FIG. S2, and FIG. S3. Bacterial DNA (16S) was found at every site analyzed in significant quantity, spanning several orders of magnitude. Archaeal (16S) and fungal (18S) DNA were also found at every site, although typically at less than one percent the level of bacterial DNA. Overall, as with ATP, samples from Eldfell showed noticeably higher concentrations of bacterial and especially archaeal DNA; however, fungal DNA did not have as obvious of a trend. The data sets for all three primers were analyzed following the same procedure described for the ATP and fluorescence microscopy data.

Like the ATP data, the bacterial and archaeal qPCR data show generally increasing variance with increasing spatial scale. (Two 1 m sites, FIM–2–1–1 and FIM–4–3–3, had only a single qPCR data point for fungal 18S DNA due to limitations of throughput in the field site, so they were excluded from this analysis and from the *F*- and *H*-tests.) The fungal qPCR data, however, shows no clear trend.

The *F* tests for bacterial and archaeal DNA show strong effects at all four spatial scales ($p < 0.001$). Fungal DNA shows similarly strong diversity at 1 m, 10 m, and 100 m, but not at the > 1 km scale. The pairwise *t*-tests show that, for relative bacterial DNA content, between a third to a half of sites are different from each other at a significance level of $\alpha = 0.05$ (43% at 1 m, 38% at 10 m, and 50% at 100 m), with Eldfell and Fimmvörðuháls being significantly different as well. Archaeal

DNA content across sites is somewhat more homogeneous, with between a quarter and a third of sites (23%, 30%, and 39%), plus the two > 1 km sites, differing. Fungal DNA content, on the other hand, is not significantly different across the > 1 km sites ($p = 0.48$), though at the smaller scales, 21% (1 m), 13% (10 m), and 25% (100 m) of sites were are different. Interestingly, the FIM–6–1 site, which has the only significantly higher level of relative cell concentration among sites, also has a clearly higher level of fungal DNA, but not of bacterial or archaeal DNA.

Consistent with these results, the non-parametric *H*-tests of variability within and between sampling sites show strong diversity ($p < 0.001$) at all four spatial scales for all three DNA types. As the throughput of the field lab limited the number of replicates that could be run with each set of primers, significant results could not be obtained from pairwise *u*-tests at the 1 m or 10 m scale. At the 100 m scale, 57% of sites are different in bacterial DNA content, 61% for archaeal DNA content, and 39% for fungal DNA content. At the > 1 km scale, as with the *t*-tests, the two sites differ in bacterial and archaeal DNA content; interestingly, the fungal DNA non-parametric results diverge from the *t*-test results, showing a marked difference between the two sites ($p < 0.001$). This may indicate that although the mean fungal DNA content does not differ between the two > 1 km sites, it is distributed differently at each.

Spatial scaling and correlation

The non-parametric Spearman's rank test was used to characterize the correlation between the mean values of each biomarker assay at sites at the 1 m, 10 m, and 100 m spatial scales. (Correlations at the > 1 km scale could not be tested, as there were only two sites.) Only a non-parametric test was performed, as the differences in unit size and scale between assay results, plus the unknown shape of any potential correlation (linear, logarithmic, asymptotic, *etc.*), make the results of a quantitative regression highly difficult to interpret.

The Spearman's test yields two parameters, the correlation coefficient $\boldsymbol{\rho}$, which describes the degree of correlation (close to +1 indicates a strongly positive correlation, close to –1 indicates a strongly negative correlation) between the ranks of the two data sets, and significance p, which describes the significance of the correlation coefficient (the lower the number, the better the proposed correlation fits the data in comparison to a non-correlated model). Scatter plots for the correlated data are shown in FIG. 7 and FIG. S4.

At the 1 m scale (TABLE S3), mean ATP content is reasonably strongly correlated with relative mean bacterial, archaeal, and fungal DNA content ($0.46 < \rho < 0.65$) with high confidence ($p < 0.005$). Relative bacterial, archaeal and fungal DNA content all have even stronger positive correlations among themselves ($0.54 < \rho < 0.81$, $p < 0.001$), with the correlation between archaea and fungi somewhat less

pronounced than the others. Cell concentration, however, is not significantly correlated with any of the other four measures, although the values for fungal DNA content are borderline ($\rho = 0.29$, $p = 0.054$).

At the 10 m scale (TABLE S4), the pattern is similar, although overall the confidence levels are slightly less strong (an effect expected due to the smaller number of sites). Cell concentration data shows only weak, non-significant positive correlations with the other four assays.

At the 100 m scale (TABLE S5), at which there are only 8 sites, correlations become even less certain; however, the general trends seen at 1 and 10 m still hold. ATP remains significantly correlated only with bacterial and fungal DNA content; the correlation between archaeal and fungal DNA is no longer significant. Once again, no relation is seen to cell concentration.

The difference in typical level of archaeal DNA content between Eldfell and Fimmvörðuháls is visible in the scatter plots. The test between archaeal and fungal qPCR data was re-run with the sets divided at the > 1 km level to see if one group was more strongly correlated than the other. There is no correlation at any spatial scale for the Eldfell data ($\rho \sim 0.5$, $p > 0.08$); like the combined data, the Fimmvörðuháls data is of borderline significance ($\rho \sim 0.5$, $p \sim 0.05$) at the 1 and 10 m scale, but not significantly correlated at the 100 m scale.

Discussion and Conclusions

Diversity and spatial scaling of sites

Five methods of assessing sample site representativeness (sample mean variance, group *F*-tests, pairwise *t*-tests, and the distribution-free rank sum *H* and *u*-tests) have been used to characterize three different biomarker assays (fluorescence microscopy, ATP, and qPCR) at four different spatial scales (1 m, 10 m, 100 m, and > 1 km) in Icelandic Mars analogue field sites. The assays were performed on samples taken under simulated robotic planetary exploration conditions with best-effort realism in numbers of sample locations, sample replicates, and repetitions of assays performed. High-level variability is summarized in TABLE 3.

At a high level, the results are generally consistent with each other and with previous work. The measured ATP concentrations span a range of 0.02 to 1.6 nM; this is at the very low end of the range of published values for other extreme environments (see Barnett (2010) and references therein), indicating that, as we hypothesized, such recent lava fields have relatively sparse microbial ecologies. The non-zero cell concentrations span a range of 5.8×10^4 to 2.0×10^6 cells per mL sample; using published estimates for the anticipated amount of ATP liberated per cell (Cowan et al., 2002), this would correspond to expected ATP ranges of 0.02 to 9.0 nM, which agrees well with our experimental results. (The qPCR results, being purely relative, cannot be compared to previous work in the same way.)

The nucleic acid stain and fluorescence microscopy assay for cell concentration shows by far the least diversity between sites at all surveyed spatial scales. The pairwise tests indicate that, under our 'typical remote sensing' definition of site homogeneity, most possible site pairs would not yield statistically distinct results under our simulated operational constraints. The group variability characterizations yield a more mixed picture; a single highly divergent site is responsible for much of the significant variability, and the more conservative non-parametric tests show borderline differences between sites at 10 m and 100 m resolutions. All tests agree that at the > 1 km scale, the two sites are not distinct. Thus, there may be local patterns of cell number diversity present at 10 m and 100 m which sampling at a finer or sparser resolution may miss.

The ATP bioluminescence assay shows the greatest amount of diversity between sites at all surveyed spatial scales on all five measures. The pairwise tests show that, at best, only half of possible site pairs would yield statistically similar results under our simulated operational criteria. In combination with the group variance characterization, these results indicate that, for this combination of biomarker and assay under this set of analogue site selection criteria, quantification of a given single sample location should not be taken as representative of likely results more than 1 m away.

The qPCR bacterial and archaeal assays yield results that are very similar to each

other; both show strong location effects at all four spatial scales, although to a somewhat lesser degree than the ATP data. The qPCR fungal data is somewhat more mixed; although all spatial scales show some diversity under the *H*-test, the percentage of distinct sites is lower, and the means of the > 1 km sites are not distinctly different. Thus, as with ATP quantification, bacterial and archaeal DNA quantification of a single sample should not be taken as representative of areas more than 1 m away under these site selection criteria. Fungal DNA content, however, appears to follow a spatial diversity distribution more like that of cell number, which has significant local variation but is nonetheless be relatively homogeneous when averaged over scales of > 1 km.

Correlation between methods

High-level correlation is summarized in TABLE 4. The final value yielded by a biomarker assay is affected by the true value at the sample site, sample collection (*e.g.*, whether the concentration of living organisms in the sample is representative of the overall site), recovery from the collected sample (*e.g.*, what fraction of living cells in the sample are delivered in the same state), losses during sample prep (*e.g.*, whether all collected cell types are effectively lysed), and how efficiently the assay itself detects the target biomarker (*e.g.*, possible matrix effects). In our case, all three assays shared the same sample collection and recovery, ATP and fluorescence microscopy both involved a sample wash to separate cells, and ATP and qPCR both used cell lysis. As shared procedural steps do not rule out different responses

between assays, all of these factors must be examined in a discussion of how these five data sets are not universally correlated with each other.

Interestingly, the assay that most directly measures biomass, cell concentration, displayed no significant correlation with ATP concentration or with bacterial, archaeal, or fungal DNA concentration. In other studies, DNA content and ATP content per cell have both been observed to vary over approximately an order of magnitude in soil samples (see Cowan et al. (2002), Zhou et al. (1996), and references therein) due to differing ratios of species, levels of metabolic activity, *etc.*; however, these same factors can also cause differences in the effectiveness of lysis of samples with similar mean content. Both of these causes would be expected to have an underlying sensitivity to environmental factors beyond those kept homogeneous in site selection here (apparent color, morphology, moisture content, and grain size), such as pH or more detailed geochemistry.

Mean ATP concentration is reasonably well correlated with relative bacterial and fungal DNA concentration and somewhat more weakly with relative archaeal DNA concentration. The overwhelming dominance of bacterial DNA in all samples means that, although there are clearly patterns in the relative percentages of the three DNA types, overall DNA content tracks bacterial DNA content very closely. ATP concentration is therefore correlated fairly well with DNA content but not with cell concentration. This gives greater weight to the hypothesis that differences in

cell type and state, and hence efficiency of lysis techniques, were present between sites.

Of the three types of DNA quantified by qPCR, bacterial DNA content correlates well with both archaeal and fungal DNA content. The correlation between archaeal and fungal DNA is inconsistent; it is not significant in the sites at Eldfell, but significant (though not particularly strong) at the 1 and 10 m scales at Fimmvörðuháls. As described above, total DNA content is dominated by bacterial DNA, so it is not surprising that levels of archaeal and fungal DNA track total DNA content; this may indicate that certain sites were more generally hospitable to life, but given the lack of correlation with cell concentration, likely also indicates that samples from certain sites allowed more efficient DNA extraction. Between archaeal and fungal DNA content, relative fungal DNA content tracks ATP concentration slightly better; as both archaeal and fungal DNA shared the same sample aliquots, preparation, and assay runs, this is more likely to indicate a true difference in site biomarker content.

Implications for assay and site selection criteria

The biomarker ATP, as measured with a bioluminescence assay, appears to be by far the most sensitive to small changes in sampling location within our simulated remote sensing definition of a consistent environment. Bacterial and archaeal DNA content appear similarly suitable for studying small-scale variation or

microenvironments, but not necessarily for drawing conclusions about large areas. Cell concentrations and fungal DNA content have significant local variation but appear relatively homogeneous when averaged over scales of > 1 km, making them better suited for wide-ranging but sparse sampling. The effects of low biomass are apparent in the low overall numbers and high variance in cell concentration. The effects of multiple potential life dispersal mechanisms and preferred habitats are apparent in the differences in quantity and distribution of the measured types of DNA content. Future work in analogue environments specifically to investigate biomarker distributions at different scales, without simulated robotic exploration constraints and with finer-resolution characterization of the physicochemical environment, appears justified.

Several commonly performed statistical tests were performed to show the range of typical results that could be expected in these planetary analogue conditions. These results highlight the need for multiple analyses to be performed in order to understand biomarker distribution and representativeness. In a robotic exploration context, sampling constraints and lack of *a priori* information on likely biomarker distribution may justify performing complementary distribution-free statistical tests; although they require a different tradeoff between sample replicates (a recommended minimum is 10) and sample location to achieve the same level of significance, the fungal qPCR data here are an example of how they may show

changes in biomarker concentration that are otherwise masked.

These findings are therefore highly relevant for future biomarker sampling expeditions, including selection of number, distribution, and location of sampling sites, selection of assays, selection of analyses, and (for DNA detection and quantification) selection of primers for sparse sampling spatial distributions. Future work will be necessary to determine how spatial scale continues to be relevant on scales smaller than 1 m or greater than 1 km, and whether the significant levels of spatial diversity co-vary with environmental variables, such as grain size, oxidation state, moisture content, and so forth.

ATP and DNA content overall correlate well with each other, meaning that using both together may increase confidence in a positive or negative result. Cell concentration, measured via nucleic acid staining, does not correlate well with either, meaning that using it with one of the others may increase the chance of a positive return or of a reliable result in the event of matrix or other environmental effects on assay yield. The inconsistent correlation between biomarker assay results, and the differing degrees of strength of correlation between those that were correlated, highlight the importance of understanding not only the likely true spatial and temporal distributions of a target biomarker, but the sample collection, preparation, and environmental matrix effects on the particular assay used to measure it.

All three biomarkers measured here (nucleic acids, ATP, and ribosomal DNA) are specific to terrestrial life. As we do not anticipate that life that has evolved elsewhere under different thermodynamic and chemical conditions would necessarily require these same constituents, we do not expect future space missions to use the set of assays described here. Our results show, at a higher level, that selecting a suite of assays designed to measure different types of biomarkers (in our case, those related to biomass, metabolic activity, and biodiversity) will be critical in understanding the potential habitability and distribution of life in extraterrestrial environments. This kind of work, performed in planetary analogue environments, is key to choosing biomarkers, instrumentation, assays, and sampling strategies that will yield the best science return within the operational constraints of life detection and robotic exploration.

Acknowledgments

The authors would like to thank Lewis and Clark Exploration Fund (ESA, MLC and EWS), Oak Ridge Associated Universities (AMS), Cranfield University (DCC), and The Open University (AHS) for their contributions to the expedition. The community of Hvolsvöllur, especially Sigurlín Sveinbjarnardóttir, the headmistress of Hvolsskóli, generously allowed the use of its classrooms as our field laboratory; the staff of the Hotel Hvolsvöllur made possible the arrangements for shipping our equipment. The authors would also like to acknowledge the

communities of Hvolsvöllur and Heimaey for allowing us to sample at their sites and the Icelandic Institute of Natural History for granting the export permits required to complete this work.

Author Disclosure Statement

All authors state that no competing financial interests exist.

References

Amador, E., Cable, M., Chaudry, N., Cullen, T., Gentry, D., Jacobsen, M., Murusekan, G., Schwieterman, E., Stevens, A., Stockton, A., Yin, C., Cullen, D., and Geppert, W. (2014) Synchronous in-field application of life-detection techniques in planetary analogue missions. *Planetary and Space Science* 106:1–10.

Bagshaw, E., Cockell, C., Magan, N., Wadham, J., Venugopalan, T., Sun, T., Mowlem, M., and Croxford, A. (2011) The microbial habitability of weathered volcanic glass inferred from continuous sensing techniques. *Astrobiology* 11:651.

Baker, G., Smith, J., and Cowan, D. (2003) Review and re-analysis of domain-specific 16s primers. *Journal of Microbiological Methods* 55:541–555.

Banerjee, S., Si, B., and Siciliano, S. (2011) Evidence of high microbial abundance and spatial dependency in three Arctic soil ecosystems. *Soil Science Society of America Journal* 75:2227–2232.

Barnett, M. (2010) *Implementation of in-field life detection and characterisation techniques in icy environments*. Ph.D. thesis, Cranfield University.

Blake, D., Vaniman, D., Achilles, C., Anderson, R., Bish, D., Bristow, T., Chen, C., Chipera, S., Crisp, J., DesăMarais, D., Downs, R.T., Farmer, J., Feldman, S., Fonda, M., Gailhanou, M., Ma, H., Ming, D.W., Morris, R.V., Sarrazin, P., Stolper, E., Treiman, A., and Yen, A. (2012) Characterization and calibration of the CheMin mineralogical instrument on Mars Science Laboratory. *Space Science Reviews* 170:341–399.

Borneman, J. and Hartin, R. (2000) PCR primers that amplify fungal rRNA genes from environmental samples. *Applied and Environmental Microbiology* 66:4356–4360.

Cousins, C. and Crawford, I. (2011) Volcano-ice interaction as a microbial habitat on Earth and Mars. *Astrobiology* 11:695–710.

Cowan, D., Russell, N., Mamais, A., and Sheppard, D. (2002) Antarctic dry valley mineral soils contain unexpectedly high levels of microbial biomass. *Extremophiles* 6:431–436.

Ettema, C.H. and Wardle, D.A. (2002) Spatial soil ecology. *Trends in ecology & evolution* 17:177–183.

Fajardo-Cavazos, P., Schuerger, A., and Nicholson, W. (2008) Persistence of biomarker ATP and ATP-generating capability in bacterial cells and spores contaminating spacecraft materials under earth conditions and in a

simulated martian environment. *Applied and Environmental Microbiology* 74:5159–5167.

Fajardo-Cavazos, P., Schuerger, A., and Nicholson, W. (2010) Exposure of DNA and *Bacillus subtilis* spores to simulated Martian environments: Use of quantitative PCR (qPCR) to measure inactivation rates of DNA to function as a template molecule. *Astrobiology* 10:403–411.

Franklin, R. and Mills, A. (2003) Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology* 44:335–346.

Green, J. and Bohannan, B. (2006) Spatial scaling of microbial biodiversity. *Trends in Ecology and Evolution* 21:501–507.

Higgins, M. and Roberge, J. (2007) Three magmatic components in the 1973 eruption of Eldfell volcano, Iceland: Evidence from plagioclase crystal size distribution (CSD) and geochemistry. *Journal of Volcanology and Geothermal Research* 161:247–260.

Kepner, R. and Pratt, J. (1994) Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present. *Microbiol Reviews* 58:603–615.

Kieft, T. (2003) Desert environments: Soil microbial communities in hot deserts. In *Encyclopedia of Environmental Microbiology*. John Wiley & Sons, Inc., pp. 1–25.

Livak, K. and Schmittgen, T. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta c(t)) method. *Methods* 25:402–408.

Mahaffy, P., Webster, C., Cabane, M., Conrad, P., Coll, P., Atreya, S., Arvey, R., Barciniak, M., Benna, M., Bleacher, L., Brinckerhoff, W., Eigenbrode, J., Carignan, D., Cascia, M., Chalmers, R., Dworkin, J., Errigo, T., Everson, P., Franz, H., Farley, R., Feng, S., Frazier, G., Freissinet, C., Glavin, D., Harpold, D., Hawk, D., Holmes, V., Johnson, C., Jones, A., Jordan, P., Kellogg, J., Lewis, J., Lyness, E., Malespin, C., Martin, D., Maurer, J., McAdam, A., McLennan, D., Nolan, T., Noriega, M., Pavlov, A., Prats, B., Raaen, E., Sheinman, O., Sheppard, D., Smith, J., Stern, J., Tan, F., Trainer, M., Ming, D., Morris, R., Jones, J., Gundersen, C., Steele, A., Wray, J., Botta, O., Leshin, L., Owen, T., Battel, S., Jakosky, B., Manning, H., Squyres, S., Navarro-González, R., McKay, C., Raulin, F., Sternberg, R., Buch, A., Sorensen, P., Kline-Schoder, R., Coscia, D., Szopa, C., Teinturier, S., Baffes, C., Feldman, J., Flesch, G., Forouhar, S., Garcia, R., Keymeulen, D., Woodward, S., Block, B., Arnett, K., Miller, R., Edmonson, C., Gorevan, S., and Mumm, E. (2012) The Sample Analysis at Mars investigation and instrument suite. *Space Science Reviews* 170:401–478.

McKay, C.P., Stoker, C.R., Glass, B.J., Davé, A.I., Davila, A.F., Heldmann,

J.L., Marinova, M.M., Fairen, A.G., Quinn, R.C., Zacny, K.A., Paulsen, G., Smith, P.H., Parro, V., Andersen, D.T., Hecht, M.H., Lacelle, D., and Pollard, W.H. (2013) The Icebreaker Life mission to Mars: A search for biomolecular evidence for life. *Astrobiology* 13:334–353.

Mendenhall, W. and Sincich, T. (1995) *Statistics for Engineering and the Sciences*. 4th edition. Prentice-Hall International.

Nadeau, J., Perreault, N., Niederberger, T., Whyte, L., Sun, H., and Leon, R. (2008) Fluorescence microscopy as a tool for *in situ* life detection. *Astrobiology* 8:859–874.

Naveed, M., Herath, L., Moldrup, P., Arthur, E., Nicolaisen, M., Norgaard, T., Ferré, T.P., and de Jonge, L.W. (2016) Spatial variability of microbial richness and diversity and relationships with soil organic carbon, texture and structure across an agricultural field. *Applied Soil Ecology* 103:44–55.

Obousy, R., Tziolas, A., Kaltsas, K., Sims, M., and Grant, W. (2000) Searching for extant life on Mars - the ATP-firefly luciferin/luciferase technique. *Journal of the British Interplanetary Society* 53:121–130.

Parnell, J., Cullen, D., Sims, M., Bowden, S., Cockell, C., Court, R., Ehrenfreund, P., Gaubert, F., Grant, W., Parro, V., Rohmer, M., Sephton, M., Stan-Lotter, H., Steele, A., Toporski, J., and Vago, J. (2007) Searching for life on Mars: selection of molecular targets for ESA's Aurora ExoMars mission. *Astrobiology* 7:578–604.

Pasternak, Z., Al-Ashhab, A., Gatica, J., Gafny, R., Avraham, S., Minz, D., Gillor, O., and Jurkevitch, E. (2013) Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions. *PloS One* 8:e69705.

Peigné, J., Vian, J.F., Cannavacciuolo, M., Bottollier, B., and Chaussod, R. (2009) Soil sampling based on field spatial variability of soil microbial indicators. *European Journal of Soil Biology* 45:488–495.

Preston, L. and Dartnell, L. (2014) Planetary habitability: lessons learned from terrestrial analogues. *International Journal of Astrobiology* 13:81–98.

Schneider, C., Rasband, W., and Eliceiri, K. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9:671–675.

Sigmundsson, F., Hreinsdóttir, S., Hooper, A., Árnadóttir, T., Pedersen, R., Roberts, M.J., Óskarsson, N., Auriac, A., Decriem, J., Einarsson, P., Geirsson, H., Hensch, M., Ófeigsson, B.G., Sturkell, E., Sveinbjörnsson, H., and Feigl, K.L. (2010) Intrusion triggering of the 2010 Eyjafjallajökull explosive eruption. *Nature* 468:426–430

Sims, M.R., Cullen, D.C., Rix, C.S., Buckley, A., Derveni, M., Evans, D., Miguel García-Con, L., Rhodes, A., Rato, C.C., Stefinovic, M., Sephton, M.A., Court, R.W., Bulloch, C., Kitchingman, I., Ali, Z., Pullan, D., Holt, J., Blake, O., Sykes, J., Samara-Ratna, P., Canali, M., Borst, G., Leeuwis, H., Prak, A., Norfini, A., Geraci, E., Tavanti, M., Brucato, J., and Holm,

N. (2012) Development status of the life marker chip instrument for ExoMars. *Planetary and Space Science* 72:129–137.

Stanley, P. (1989) A review of bioluminescent ATP techniques in rapid microbiology. *Journal of Bioluminescence and Chemiluminescence* 4:375–380.

Treiman, A.H., Bish, D.L., Vaniman, D.T., Chipera, S.J., Blake, D.F., Ming, D.W., Morris, R.V., Bristow, T.F., Morrison, S.M., Baker, M.B., Rampe, E.B., Downs, R.T., Filiberto, J., Glazner, A.F., Gellert, R., Thompson, L.M., Schmidt, M.E., Le Deit, L., Wiens, R.C., McAdam, A.C., Achilles, C.N., Edgett, K.S., Farmer, J.D., Fendrich, K.V., Grotzinger, J.P., Gupta, S., Morookian, J.M., Newcombe, M.E., Rice, M.S., Spray, J.G., Stolper, E.M., Sumner, D.Y., Vasavada, A.R., and Yen, A.S. (2016) Mineralogy, provenance, and diagenesis of a potassic basaltic sandstone on Mars: CheMin X-ray diffraction of the Windjana sample (Kimberley area, Gale Crater). *Journal of Geophysical Research: Planets* 121:75–106.

Warner, N. and Farmer, J. (2010) Subglacial hydrothermal alteration minerals in jökulhlaup deposits of southern Iceland, with implications for detecting past or present habitable environments on Mars. *Astrobiology* 10:523–547.

Yu, Y., Lee, C., Kim, J., and Hwang, S. (2005) Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnology and Bioengineering* 89:670–

679.

Zhou, J., Bruns, M., and Tiedje, J. (1996) DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* 62:316–322.

Figure Captions

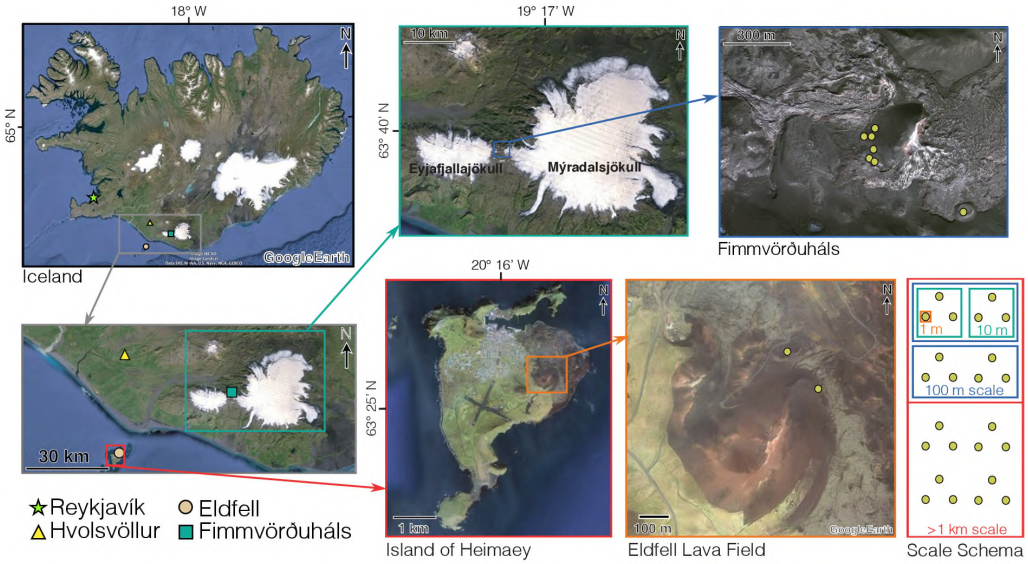


FIG. 1: The two major field sites: Eldfell, a decades-old eruption site on the island of Heimaey, and Fimmvörðuháls, a region between glaciers with access to both very recent and much older lava fields. At each major site, triplicate samples were taken at each point in a 1 m triangle; this pattern was repeated at 2-3 locations 10 m apart, and then that nested pattern repeated 2-3 times at locations 100 m apart.

FIG. 2: Representative photographs taken at each field site. Sites were selected to be homogeneous (*e.g.*, color, morphology, and grain size) at coarse remote sensing resolution. **(a,b)** Eldfell sampling locations. **(c)** Preparations for field sampling near the Eldfell scoria cone. **(d)** Heimaey, looking towards Eldfell. Note the beginnings of vegetation growth and anthropogenic tracks on this side of the cone. **(e)** The cinder cone in the Fimmvörðuháls field. **(f,g)** Fimmvörðuháls sampling locations. **(h**) Preparations for field sampling at Fimmvörðuháls.

FIG. 3: The field lab during processing of Fimmvörðuháls samples: **(a)** DNA

extraction; **(b)** qPCR; **(c)** nucleic acid staining; **(d)** ATP luminescence; **(e)** shared sterile consumables. Stations not shown: sample prep (grinding) room, dark room for microscopy.

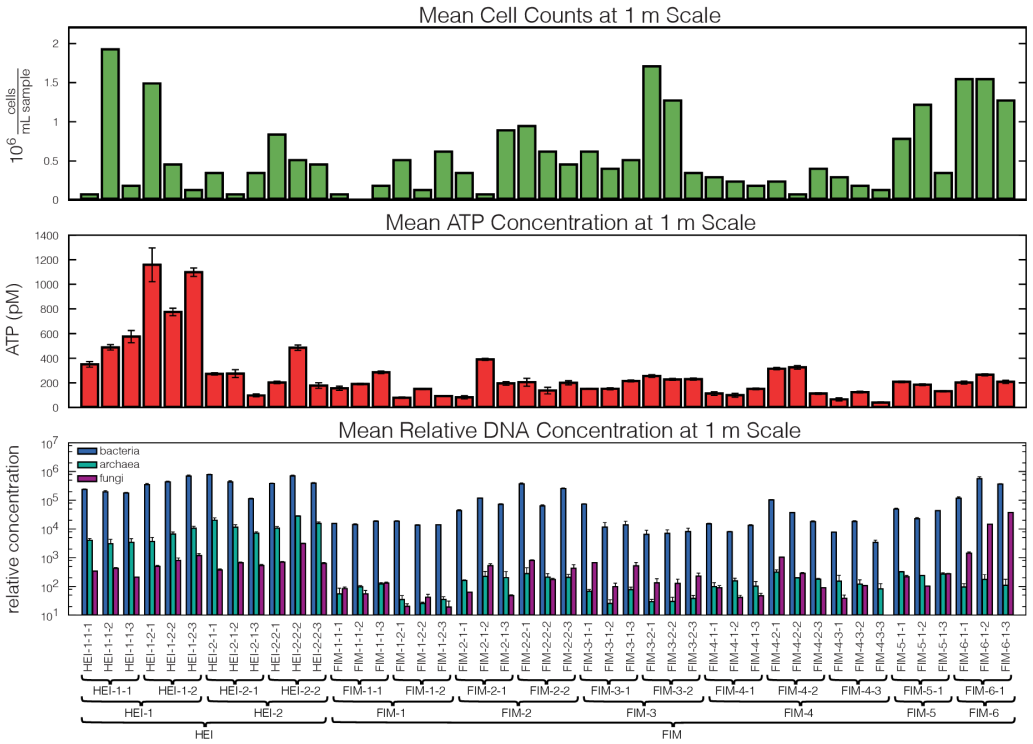
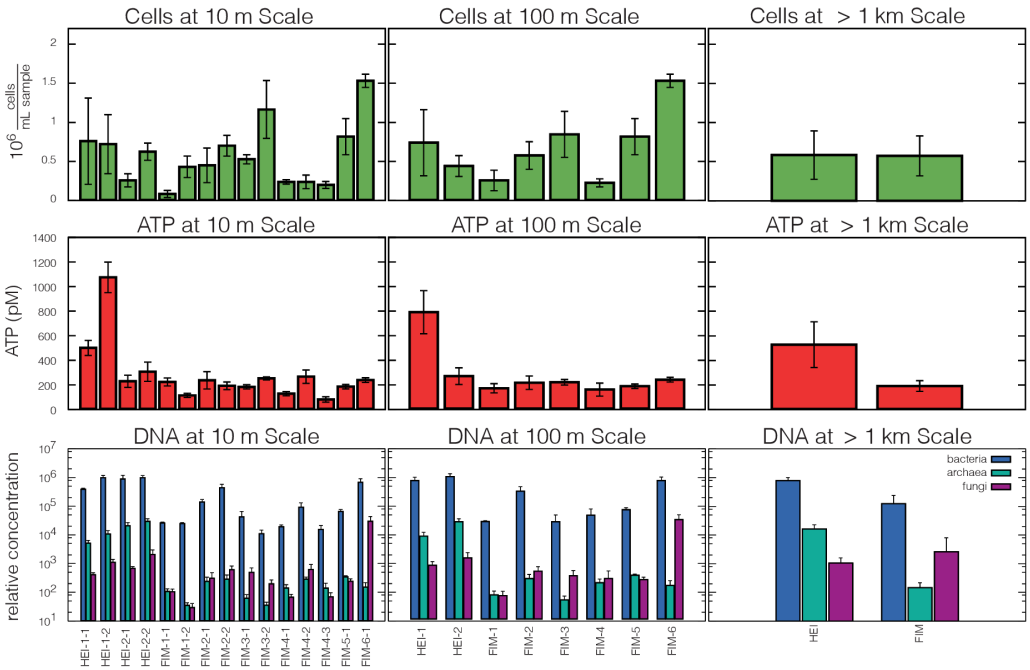


FIG. 4: Mean values for all assays at the 1 m scale. ATP values are means of four replicates. Cell concentration values are single samples. 16S/18S DNA content values are means of two replicates, with the exception of fungal 18S data at FIM-2-1-1 and FIM-4-3-3, which each had only a single replicate; scale is relative to minimum detection threshold. Bars represent one standard deviation.

FIG. 5: Mean values for all assays at the 10 m, 100 m, and > 1 km scales. Data is



plotted as in FIG. 4.

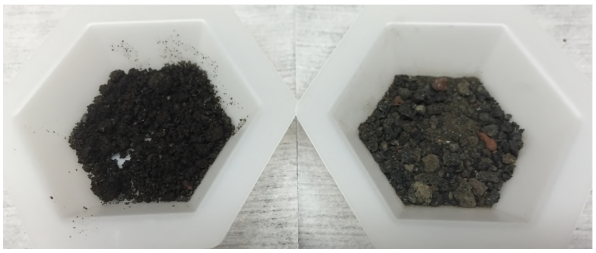


FIG. 6: Samples from FIM-4-1 (left) and FIM-6-1 (right); the FIM-6-1 sample has a notably higher percentage of reddish (most likely oxidized) particles.

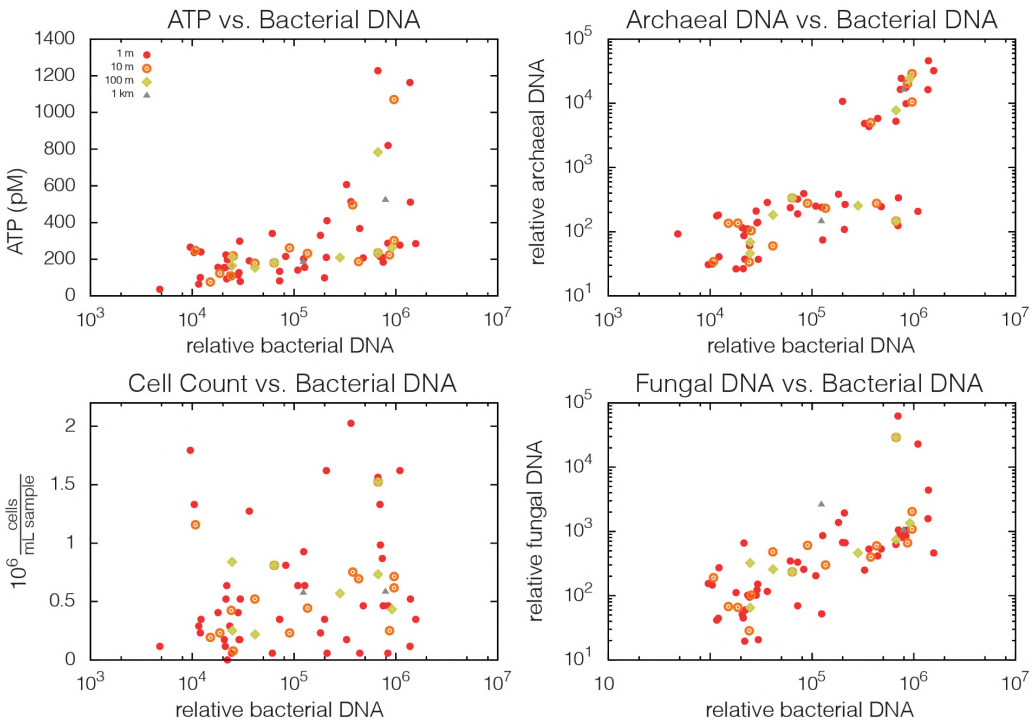


FIG. 7: Mean relative 16S bacterial DNA concentrations correlated with mean ATP concentration, mean cell concentrations, mean relative 16S archaeal DNA content, and mean relative 18S fungal DNA content at all sampling sites and scales.

Tables

TABLE 1: Primers used in qPCR assays. Bacterial primer names are those published by Baker et al. (2003), fungal primers those published by Borneman and Hartin (2000), and archaeal primers those published by Yu et al. (2005).

Target		Forward Primer	Reverse Primer	Extension Temp.
bacteria	16S DNA	E85	E533R	54 ° C
archaea	16S DNA	ARC787F	ARC1059R	54 ° C
fungi	18S DNA	nu-SSU-0817	nu-SSU-1196	54 ° C

TABLE 2: Mean variances of the results of each life detection technique at all four spatial scales. (The fluorescence microscopy data did not have sufficient replicates for complete analysis at the 1 m scale; two locations in the qPCR fungal data at 1 m, FIM-2-1-1 and FIM-4-3-3, were excluded for the same reason.)

	FM $\left(\frac{\text{cells}}{\text{mL sample}}\right)^2$	**ATP** $\left(\frac{\text{mol}}{\text{L}}\right)^2$	**qPCR (unitless)**		
Scale			bacteria	archaea	fungi
1 m		3.19×10^{-21}	9.93×10^{9}	1.09×10^{7}	6.77×10^{4}
10 m	2.04×10^{11}	1.07×10^{-20}	6.22×10^{10}	2.69×10^{7}	5.07×10^{7}
100 m	1.97×10^{11}	2.18×10^{-20}	8.11×10^{10}	2.66×10^{7}	9.48×10^{7}
> 1 km	3.21×10^{11}	7.27×10^{-20}	1.30×10^{11}	8.75×10^{7}	5.79×10^{7}

TABLE 3: High-level variability as measured by all four quantitative tests at all four scales. 'Low' indicates $p > 0.1$ or less than 10% difference, 'mid' $0.1 > p > 0.01$ (excluding outliers) or less than 50% difference, and 'high' indicates $p < 0.01$ or at least 50% difference.

		FM	**ATP**	**qPCR**		
				bacteria	archaea	fungi
F	1 m	n/a	high	high	high	high
	10 m	mid	high	high	high	high
	100 m	mid	high	high	high	high
	> 1 km	low	high	high	high	low
H	1 m	n/a	n/a	high	high	high
	10 m	low	high	high	high	high
	100 m	mid	high	high	high	high
	> 1 km	low	high	high	high	high
t	1 m	n/a	mid	mid	mid	mid
	10 m	mid	mid	mid	mid	mid
	100 m	mid	mid	high	mid	mid
	> 1 km	low	high	high	high	low
u	1 m	n/a	n/a	n/a	n/a	n/a
	10 m	n/a	high	n/a	n/a	n/a
	100 m	n/a	high	high	high	high
	> 1 km	low	high	high	high	high

TABLE 4: High-level correlation as measured by all four quantitative tests at the 1 m scale. 'Low' indicates $p > 0.1$, 'mid' $0.1 > p > 0.01$ (excluding outliers), and 'high' indicates $p < 0.01$.

		FM	**ATP**		**qPCR**	
				bacteria	archaea	fungi
FM			low	low	low	mid
ATP				mid	mid	high
	bacteria				high	high
qPCR	archaea					high
	fungi					

Supplementary Figure Captions

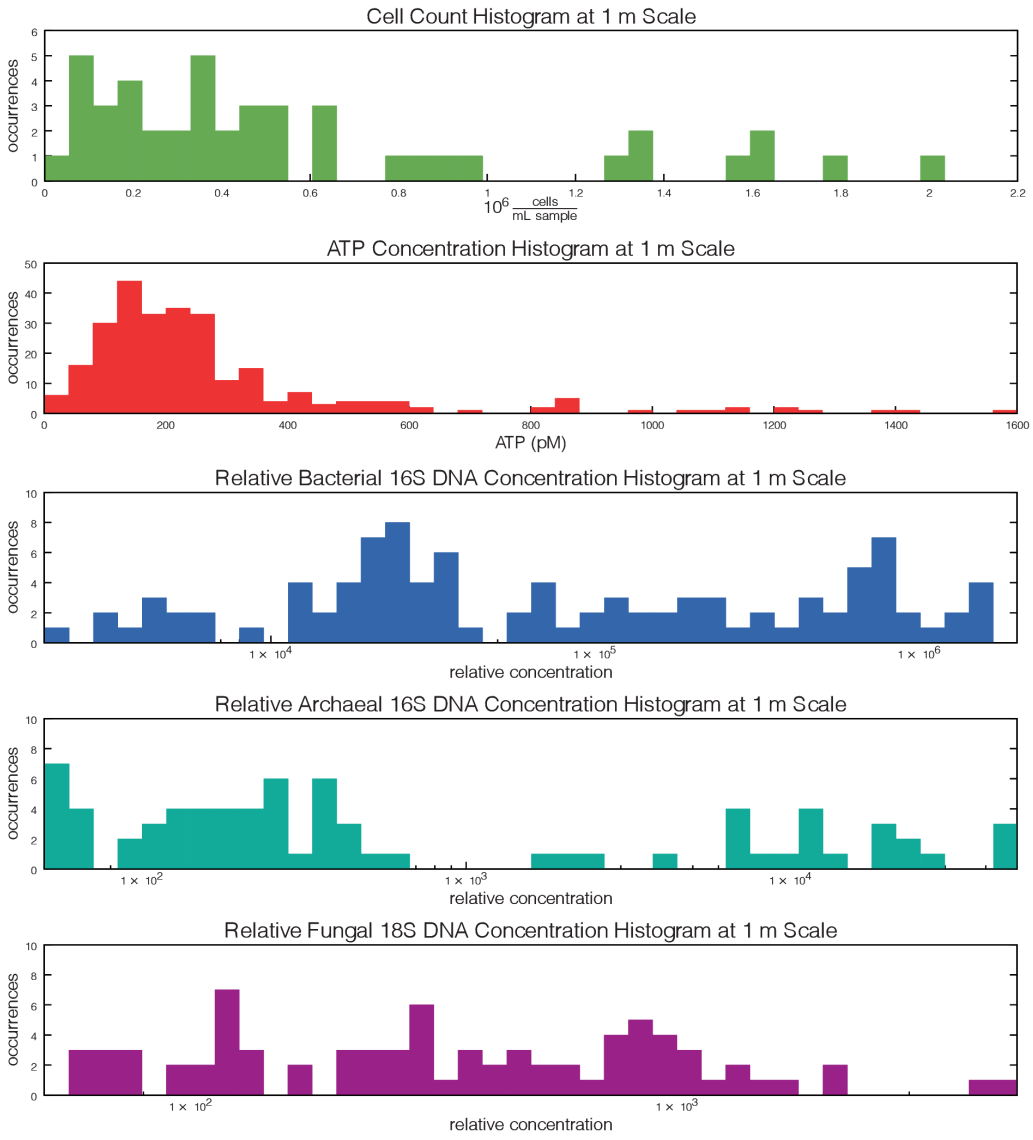
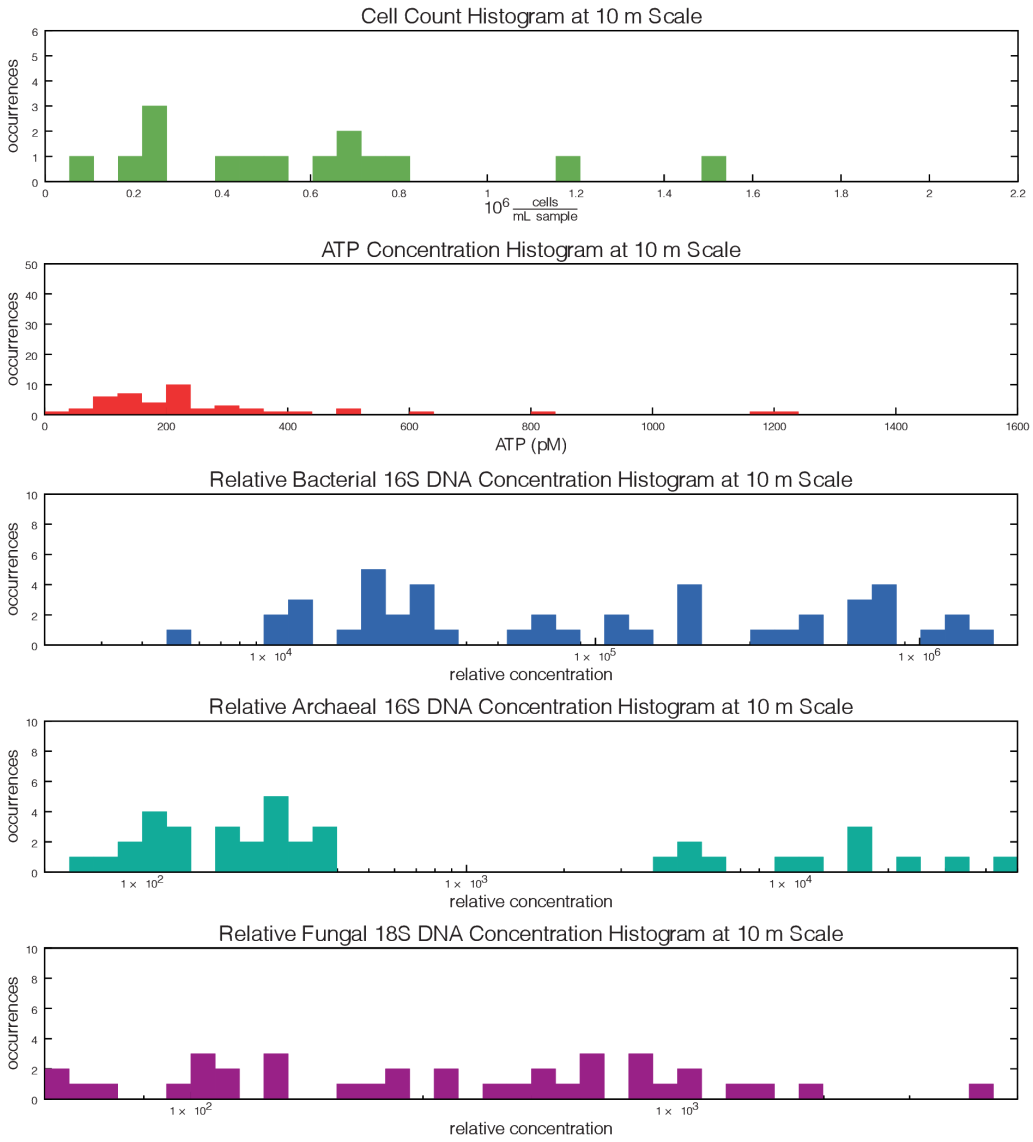


FIG. S1: Histograms showing number of occurrences of values within each range interval for all replicates measured with all assays at the 1 m scale: ATP concentration, mean cell concentrations per mL sample, relative 16S bacterial DNA

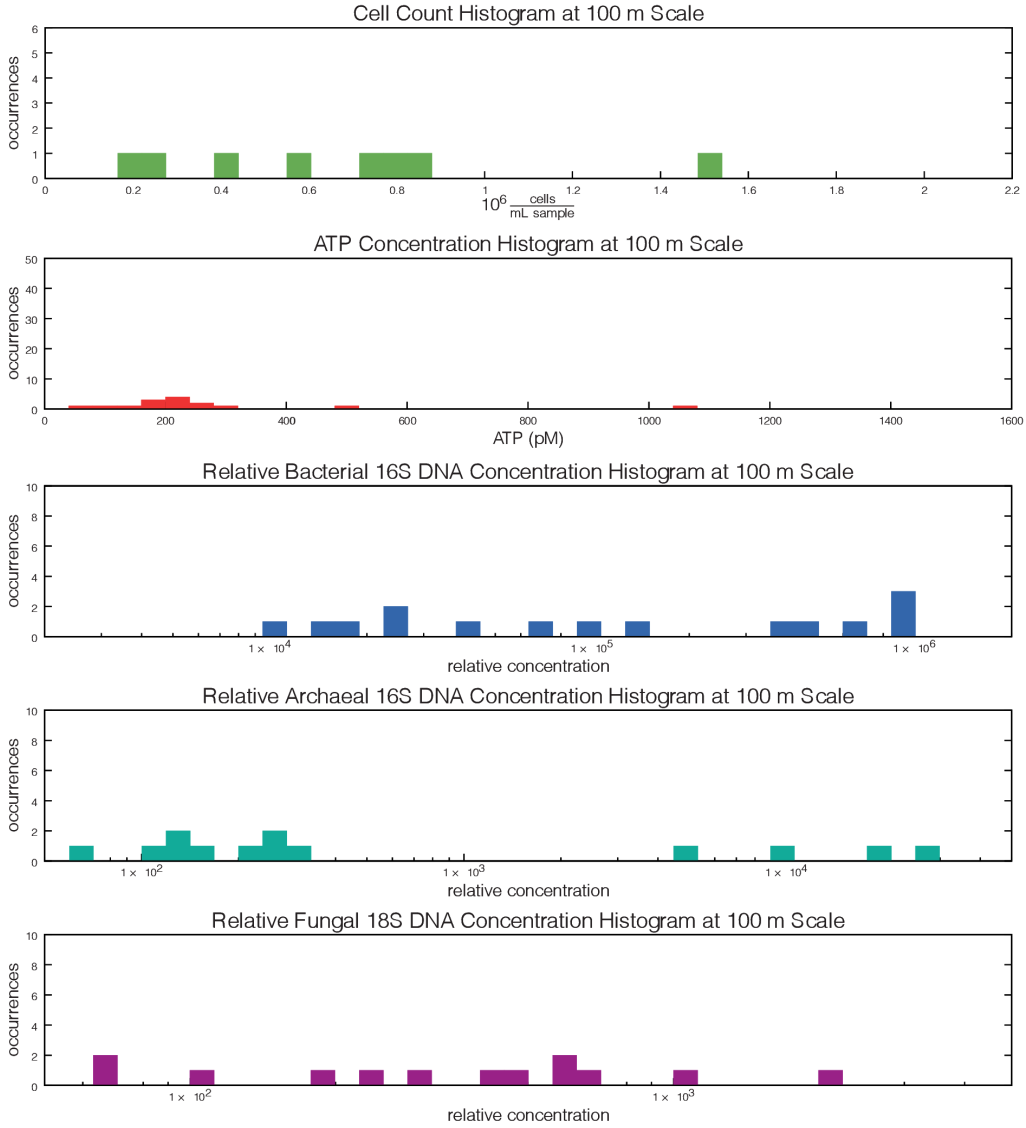
content, relative 16S archaeal DNA content, and relative 18S fungal DNA content. All histograms have 40 bins spanning the maximum and minimum data values; bins are log-scaled for DNA content data.

FIG. S2: Histograms showing number of occurrences of values within each range interval for mean values measured with all assays at the 10 m scale. Data plotted as



in FIG. S1.

FIG. S3: Histograms showing number of occurrences of values within each range interval for mean values measured with all assays at the 100 m scale. Data plotted


Cell Count Histogram at 100 m Scale
occurrences
$10^6 \frac{\text{cells}}{\text{mL sample}}$
ATP Concentration Histogram at 100 m Scale
occurrences
ATP (pM)
Relative Bacterial 16S DNA Concentration Histogram at 100 m Scale
occurrences
relative concentration
Relative Archaeal 16S DNA Concentration Histogram at 100 m Scale
occurrences
relative concentration
Relative Fungal 18S DNA Concentration Histogram at 100 m Scale
occurrences
relative concentration

as in FIG. S1.

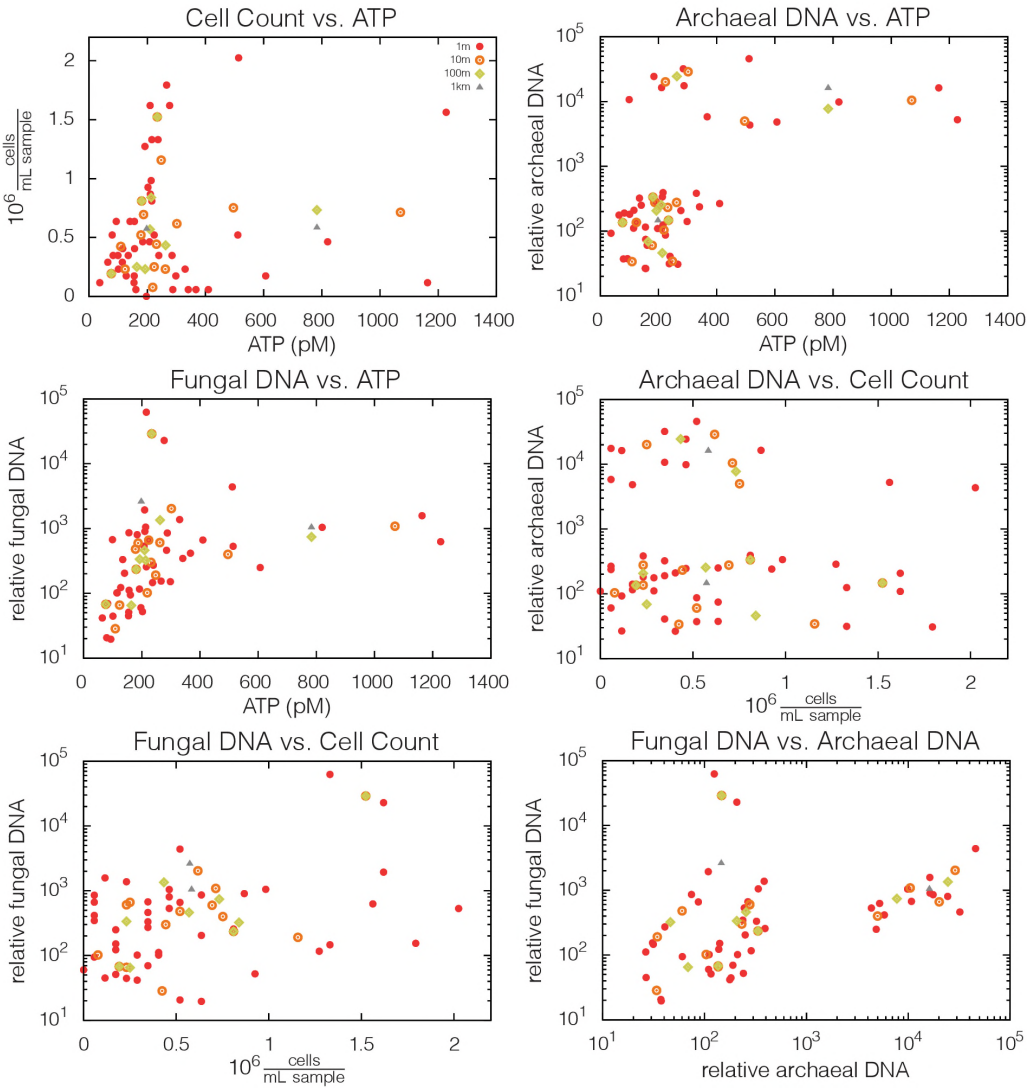


FIG. S4: Scatter plots showing correlation between data sets not presented in FIG. 7: mean cell concentrations per mL sample, mean ATP concentration, mean relative 16S archaeal DNA content, and mean relative 18S fungal DNA content at all sampling sites and scales.

Supplementary Tables

TABLE S1: Results, reported as *p (F, df_{test}, $df_{complete}$)*, for *F*-tests conducted for the results of each life detection technique at all four spatial scales. (The fluorescence microscopy data did not have sufficient replicates for complete analysis at the 1 m scale.)

	FM	**ATP**	**qPCR**		
Scale			bacteria	archaea	fungi
1 m		< 0.001 (117.78, 44, 225)	< 0.001 (46.34, 44, 55)	< 0.001 (20.77, 44, 55)	< 0.001 (20.77, 44, 53)
10 m	0.031 (2.25, 14, 30)	< 0.001 (97.28, 14, 255)	< 0.001 (16.11, 14, 85)	< 0.001 (20.01, 14, 85)	< 0.001 (20.01, 14, 83)
100 m	0.003 (3.91, 7, 37)	< 0.001 (66.00, 7, 262)	< 0.001 (24.21, 7, 92)	< 0.001 (35.65, 7, 92)	< 0.001 (35.65, 7, 90)
> 1 km	0.956 (< 0.01, 1, 43)	< 0.001 (142.15, 1, 268)	< 0.001 (89.42, 1, 98)	< 0.001 (113.21, 1, 98)	0.477 (113.21, 1, 96)

TABLE S2: Results, reported as *p (H, df_{test}, $df_{complete}$)*, for *H*-tests conducted for the results of each life detection technique at all four spatial scales. (The fluorescence microscopy data did not have sufficient replicates for complete analysis at the 1 m scale.)

	FM	**ATP**	**qPCR**		
Scale			bacteria	archaea	fungi
1 m		< 0.001 (252.03, 44, 225)	< 0.001 (94.34, 44, 55)	< 0.001 (81.14, 44, 55)	< 0.001 (89.12, 44, 53)
10 m	0.050 (23.67, 14, 30)	< 0.001 (166.44, 14, 255)	< 0.001 (85.22, 14, 85)	< 0.001 (77.00, 14, 85)	< 0.001 (66.85, 14, 83)
100 m	0.016 (17.17, 7, 37)	< 0.001 (118.10, 7, 262)	< 0.001 (76.95, 7, 92)	< 0.001 (73.00, 7, 92)	< 0.001 (52.48, 7, 90)
> 1 km	0.868 (0.03, 1, 43)	< 0.001 (68.70, 1, 268)	< 0.001 (42.79, 1, 98)	< 0.001 (54.18, 1, 98)	< 0.001 (19.65, 1, 96)

TABLE S3: Results, reported as ρ *(p)*, for Spearman's rank correlation tests conducted for the results of each life detection technique at the 1 m scale. Results significant at a level of $\alpha = 0.05$ are shown in bold.

		FM	**ATP**	**qPCR**		
				bacteria	archaea	fungi
FM			0.069(0.651)	0.177(0.246)	-0.027(0.859)	0.289(0.054)
ATP				**0.563**(**<0.001**)	**0.469**(**0.001**)	**0.644**(**<0.001**)
qPCR	bacteria				**0.803**(**<0.001**)	**0.787**(**<0.001**)
	archaea					**0.547**(**<0.001**)
	fungi					

TABLE S4: Results, reported as ρ *(p)*, for Spearman's rank correlation tests conducted for the results of each life detection technique at the 10 m scale. Results significant at a level of $\alpha = 0.05$ are shown in bold.

		FM	**ATP**		**qPCR**					
					bacteria		archaea		fungi	
FM			0.456	(0.088)	0.3276	(0.234)	0.216	(0.439)	0.431	(0.109)
ATP					**0.629**	**(0.012)**	**0.579**	**(0.024)**	**0.679**	**(0.005)**
qPCR	bacteria						**0.843**	(**<0.001**)	**0.879**	(**<0.001**)
	archaea								**0.679**	(**0.005**)
	fungi									

TABLE S5: Results, reported as ρ *(p)*, for Spearman's rank correlation tests conducted for the results of each life detection technique at the 100 m scale. Results significant at a level of $\alpha = 0.05$ are shown in bold.

		FM	**ATP**	**qPCR**		
				bacteria	archaea	fungi
FM			0.500(0.207)	0.071(0.867)	-0.190(0.651)	0.429(0.289)
ATP				**0.714**(**0.0465**)	0.452(0.26)	**0.810**(**0.0149**)
qPCR	bacteria				**0.833**(**0.0102**)	**0.762**(**0.028**)
	archaea					0.357(0.385)
	fungi					

2017-10-01

Correlations between life-detection techniques and implications for sampling site selection in planetary analog missions

Gentry, Diana M.

Mary Ann Liebert, Inc.

Gentry DM, Amador ES, Cable ML et al. (2017) Correlations between life-detection techniques and implications for sampling site selection in planetary analog missions. Astrobiology, Volume 17, Issue 12, October 2017, pp.1009-1021

https://doi.org/10.1089/ast.2016.1575